

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 991 216 A2

(12)

## EUROPEAN PATENT APPLICATION

(43) Date of publication:  
05.04.2000 Bulletin 2000/14

(51) Int. Cl.<sup>7</sup>: H04J 3/06

(21) Application number: 99111373.9

(22) Date of filing: 10.06.1999

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE  
Designated Extension States:  
AL LT LV MK RO SI

(72) Inventor: Eidson, John C.  
Palo Alto, CA 94303 (US)

(74) Representative:  
Schoppe, Fritz, Dipl.-Ing.  
Schoppe, Zimmermann & Stöckeler  
Patentanwälte  
Postfach 71 08 67  
81458 München (DE)

(30) Priority: 29.09.1998 US 162609

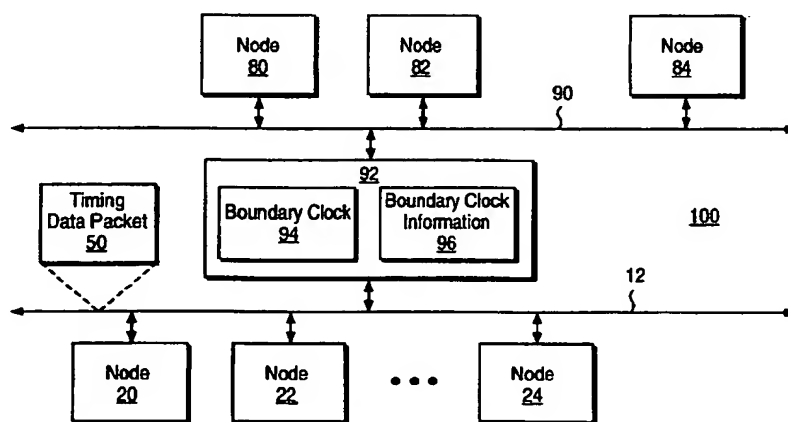
(71) Applicant:  
Hewlett-Packard Company  
Palo Alto, California 94304 (US)

## (54) Organization of time synchronization in a distributed system

(57) A distributed system with mechanisms for automatic selection of master and slave clocks to be used for clock synchronization. The distributed system includes a set of nodes (20-24), including a first node and a second node, each having a local clock (30-34) and a set of information (40-44) pertaining to the local clock (30-34). The first node transfers a packet (50) on a communication link (12) that carries the information (52). The second node receives the packet (50) on the communication link (12) and determines whether the local clock of the second node is a master clock that

synchronizes a time value in the local clock of the first node or a slave clock that synchronizes to a time value from the local clock in the first node by comparing the information (52) in the packet (50) to the information pertaining to the local clock in the second node. Automatic selection of master and slave clocks in boundary nodes (92) is provided along with mechanisms for determining clock synchronization delays and mechanisms for reporting jitter associated with communication devices (14).

FIG. 3



EP 0 991 216 A2

## Description

### BACKGROUND OF THE INVENTION

#### Field of Invention

[0001] The present invention pertains to the field of distributed systems. More particularly, this invention relates to organization of time synchronization in a distributed system.

#### Art Background

[0002] Distributed control systems are commonly arranged as a collection of nodes which are interconnected via one or more network communication links. These network communication links may be packetized links such as Ethernet or one or more of a variety of other packetized links that are adapted to distributed control system applications.

[0003] Distributed control systems commonly benefit from precise control of the timing at the distributed nodes. U.S. Patent No. 5,566,180 of Eidson et. al. teaches a method and apparatus for providing precise control of timing in distributed nodes by synchronizing the local clocks in the distributed nodes. The techniques taught by Eidson et al. use a master clock/slave clock synchronization protocol. It would be desirable to provide a technique that allows the distributed nodes to automatically select which are the master nodes and which are the slave nodes at startup and as new nodes are installed or removed or temporarily disabled. This would reduce the administrative overhead needed to maintain the distributed system. It is also desirable that such automatic selection of master and slave nodes minimally impact the communication in the distributed system.

### SUMMARY OF THE INVENTION

[0004] A distributed system is disclosed with mechanisms for automatic selection of the master and slave clocks used for clock synchronization. The distributed system includes a set of nodes, including a first node and a second node, each having a local clock and a set of information pertaining to the local clock. The first node transfers a packet on a communication link that carries the information pertaining to the local clock in the first node. The second node receives the packet on the communication link and determines whether the local clock in the second node is to be a master clock that synchronizes a time value in the local clock of the first node or a slave clock that synchronizes to a time value from the local clock in the first node. The second node makes this determination by comparing the information in the packet to the information pertaining to the local clock in the second node.

[0005] Also disclosed is automatic selection of master

and slave clocks in boundary nodes that couple together different subnets of nodes. In addition, mechanisms for determining clock synchronization delays and mechanisms for reporting jitter associated with communication devices are disclosed.

[0006] The present techniques enable a set of substantially similar code to execute in each of the nodes with some modification for boundary nodes which are slightly different due to their multi-sided structure. These techniques also provide for minimal network traffic and tolerate and correct for the appearance or disappearance of individual nodes in the system.

[0007] Other features and advantages of the present invention will be apparent from the detailed description that follows.

### BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The present invention is described with respect to particular exemplary embodiments thereof and reference is accordingly made to the drawings in which:

Figure 1 shows a distributed system which embodies techniques for automatic selection of master and slave clocks for use with clock synchronization;

Figure 2 shows one embodiment of a method for determining which of the nodes is to function as the master clock for the distributed system;

Figure 3 shows a distributed system that includes sub-nets coupled to a boundary node;

Figure 4 shows an arrangement for determining a delay associated with the transfer of timing data packets between a pair of nodes;

Figure 5 shows a communication device which includes mechanisms for determining and reporting the amount of jitter introduced in the communication device.

### DETAILED DESCRIPTION

[0009] Figure 1 shows a distributed system 10 which embodies techniques for automatic selection of the master and slave clocks for use in clock synchronization among a set of nodes 20-24 coupled to a communication link 12. The nodes 20-24 include a set of local clocks 30-34, respectively. The local clocks 30-34 keep local time for the respective nodes 20-24.

[0010] Each of the nodes 20-24 includes elements that support a predetermined synchronization protocol for synchronizing the local clocks 30-34. According to the predetermined synchronization protocol, one of the nodes 20-24 functions as a master clock for the distributed system 10 and the remaining ones of the nodes 20-24 function as slave clocks. The one of the nodes

20-24 that is the master clock periodically generates a timing data packet and transfers it via the communication link 12. Each one of the nodes 20-24 that is a slave clock receives the timing data packet and in response latches a local time value from its local clock 30-34. The node 20-24 that is the master clock generates a follow up packet for each timing data packet and transfers it via the communication link 12. Each follow up packet includes a time-stamp. Each node that is a slave clock receives the follow up packet and compares the time-stamp to the latched local time value. Each node that is a slave clock uses the difference between the time-stamp and the latched local time value to adjust its local clock 30-34.

[0011] In one embodiment, the predetermined synchronization protocol and related mechanisms implemented in the nodes 20-24 are those described in U.S. Patent No. 5,566,180. For example, each of the nodes 20-24 may include circuitry for latching a local time value when a timing data packet is recognized. Each of the slave clocks may include circuitry for adjusting its respective local clock 30-34 based upon computations of the sending and receiving time of the timing data packets which are transferred over the communication link 12. The adjustment of a stored time value may be accomplished by implementing each local clock 30-34 as a counter driven by an oscillator with sufficient stability. The least significant few bits of the counter may be implemented as an adder so that the increment on oscillator periods may be occasionally increased or decreased to effectively speed up or slow down a local clock in accordance with the results of the computation.

[0012] The nodes 20-24 may be any type of node in the distributed system 10. For example, any one or more of the nodes 20-24 may be a sensor node or an actuator node or an application controller node or a combination of these in a distributed control system. Any one or more of the nodes 20-24 may be a computer system such as a personal computer with the processor being used to calculate clock adjustment parameters.

[0013] The communication link 12 may be implemented with one or more of a variety of communication mechanisms. In one embodiment, the communication link 12 is an Ethernet communication network. In another embodiment, the communication link 12 is a LonTalk field-level control bus which is specialized for the process control environment. In other embodiments, the communication link 12 may be implemented with time division multiple access (TDMA) or token ring protocols to name only a few possibilities.

[0014] The nodes 20-24 implement methods for determining which of the nodes 20-24 is to function as the master clock for the distributed system 10. These methods make use of sets of local clock information 40-44 contained in the nodes 20-24, respectively. The local clock information 40-44 provides information pertaining to the local clocks 30-32, respectively.

[0015] Each of the nodes 20-24 initially assumes that

it is the master clock and begins transferring timing data packets via the communication link 12. Each timing data packet carries the local clock information 40-44 from the originating node 20-24. In response to a timing data packet, the nodes 20-24 compare the local clock information, contained in the timing data packet to the corresponding local clock information 40-44 to determine whether a "better" master clock is transmitting. If a better master clock is detected by a node, then that node assumes the role of a slave clock and stops generating timing data packets.

[0016] For example, at startup the node 20 assumes that it is the master clock of the distributed system 10. This may be during the initialization of the distributed system 10 or when the node 20 is installed or restarts after removal or failure or loss of communication, etc. As the master clock, the node 20 periodically generates a timing data packet, such as a timing data packet 50, and transfers it via the communication link 12. The timing data packet 50 carries a set of master clock information 52 obtained from the local clock information 40. The nodes 22-24 receive the timing data packet 50 and compare the master clock information 52 to the local clock information 42-44, respectively, to determine whether the node 20 has a better clock for the purposes of selecting a master clock for the distributed system 10.

[0017] The local clock information 40 may include an indication of whether the local clock 30 is driven by a standard time source. A standard time source may be a global positioning system (GPS) receiver or radio broadcast time source such as WWV or an atomic clock associated with the node 20. The local clock information 40 may include an indication of whether the local clock 30 appears to be functional and within specification.

[0018] The local clock information 40 may include an indication of whether the local clock 30 is traceable to a standard clock such as a GPS receiver. For example, the local clock 30 may be a very accurate clock such as an atomic clock that was at one time synchronized to a standard time source. If so, the accuracy of the local clock 30 would allow it to maintain time that is traceable to a standard clock with a high degree of accuracy even after communication with the standard clock is lost.

[0019] The local clock information 40 may include an estimate of the accuracy and/or jitter associated with the local clock 30. The local clock information 40 may include a unique identifier associated with the local clock 30. The local clock information 40 may include an indication of whether the local clock 30 is a boundary clock in a boundary node that links together subnets.

[0020] Figure 2 shows one embodiment of a method for determining which of the nodes 20-24 is to function as the master clock for the distributed system 10. The method steps shown are implemented in each of the nodes 20-24 and are discussed for the node 24.

[0021] At step 60, the node 24 waits for the detection of a timing data packet (TDP) on the communication link 12. At step 62, if no TDP is detected within a predeter-

mined timeout interval then control proceeds to step 64. Otherwise, control proceeds to step 72.

**[0022]** At step 64, the node 24 assumes the role of master clock, if not already in the role of the master clock, and begins issuing TDPs over the communication link 12 at regular intervals. Each TDP issued by the node 24 carries the local clock information 44 as master clock information to be used by receiving nodes 20-22 in evaluating which is the better master clock. The node 24 then proceeds back to step 60 to detect TDPs from one of the nodes 20-22 which may be a better master clock.

**[0023]** At step 72, the node 24 determines whether the TDP detected at step 60 originated with a new master clock. A new master clock is a master clock that has previously issued fewer than a predetermined number of TDPs which were detected by the node 24. The node 24 may determine whether TDPs were previously received from particular master clocks by logging the UIDs contained in the TDPs and recording counts associated with the UIDs logged. If it is a new master clock at step 72 then the UID of TDP detected at step 60 is logged and a count associated with that UID is incremented. The TDP is ignored and control proceeds back to step 60 to detect another TDP. If it is not a new master clock then control proceeds to step 66.

**[0024]** At step 66, the node 24 determines whether the local clock 44 or the local clock of the node that originated the TDP detected at step 60 is the best master clock. Assume in the following that the TDP detected at step 60 is the timing data packet 50. At step 66, the node 24 determines the best master clock by comparing the master clock information 52 to the local clock information 44 using a prioritized set of rules.

**[0025]** In an example embodiment, the prioritized set of rules include a first rule which involves a determination of whether the master clock information 52 or the local clock information 44 indicate a standard time source. If the master clock information 52 indicates a standard time source and the local clock information 44 does not then the local clock 34 is not the best master clock. If the local clock information 44 indicates a standard time source and the master clock information 52 does not then the local clock 34 is the best master clock at step 66. If both the master clock information 52 and the local clock information 44 indicate a standard time source then the best standard time source is selected at step 66 according to a predetermined stratum of standard clocks. For example, a GPS time source may outrank a WWV time source, etc. If the first rule fails to resolve which is the best master clock then a second rule is used at step 66.

**[0026]** In the example embodiment, the second rule involves a determination of whether the master clock information 52 or the local clock information 44 indicate a traceable time source. If the master clock information 52 indicates a traceable time source and the local clock information 44 does not then the local clock 34 is not the

best master clock. If the local clock information 44 indicates a traceable time source and the master clock information 52 does not then the local clock 34 is the best master clock at step 66. If the second rule fails to resolve which is the best master clock then a third rule is used at step 66.

**[0027]** In the example embodiment, the third rule involves a determination of whether the master clock information 52 or the local clock information 44 indicate a boundary clock. If the master clock information 52 indicates a boundary clock and the local clock information 44 does not then the local clock 34 is not the best master clock. If the local clock information 44 indicates a boundary clock and the master clock information 52 does not then the local clock 34 is the best master clock at step 66. If the third rule fails to resolve which is the best master clock then a fourth rule is used at step 66.

**[0028]** In the example embodiment, the fourth rule involves a determination of whether the master clock information 52 or the local clock information 44 indicates a better clock accuracy. If the master clock information 52 indicates a better accuracy than the local clock information 44 then the local clock 34 is not the best master clock. If the local clock information 44 indicates a better accuracy than the master clock information 52 then the local clock 34 is the best master clock at step 66. If the fourth rule fails to resolve which is the best master clock then a fifth rule is used at step 66.

**[0029]** In the example embodiment, the fifth rule is a tie breaker. For example, the clock with the lowest UID indicated in the master clock information 52 and the local clock information 44 may be selected as the best master clock at step 66.

**[0030]** Figure 3 shows a distributed system 100 that includes one sub-net having a set of nodes 80-84 coupled to a communication link 90 and another subnet that includes the nodes 20-24 coupled to the communication link 12. The distributed system 100 includes a boundary node 92 coupled between the communication links 90-12. A first side of the boundary node 92 corresponds to the subnet of the communication link 12 and a second side corresponds to the subnet of the communication link 90. A boundary node in general has n sides for linking to n subnets.

**[0031]** The boundary node 92 includes a boundary clock 94 which may assume the role of master clock on the first or the second side of the boundary node 92 or on both the first and second sides. The boundary clock 94 may also assume the role of a slave clock on either the first or the second sides of the boundary node 92. A boundary clock in a boundary node in general assumes the role of a slave clock on at most one of its sides.

**[0032]** The boundary node 92 implements the method steps 60-72 and participates in the selection of the best master clock on both its first and second sides independently. The boundary node 92 transfers timing data packets that carry a set of boundary clock information 96 via the communication link 12 when determining

which is the best master clock on the first side. Similarly, the boundary node 92 transfers timing data packets that carry the boundary clock information 96 via the communication link 90 when determining which is the best master clock on the second side.

**[0033]** The boundary node 92 includes circuitry for adjusting the boundary clock 94 from either the first or second side depending on which side, if any, the boundary clock 94 assumes the role of slave clock. In addition, the boundary node 92 includes circuitry for distributing time values using timing data packets and follow up packets to the first and/or second sides depending on which sides the boundary clock 94 assumes the role of master clock.

**[0034]** If the boundary clock 94 assumes the role of a slave clock on one of the sides of the boundary node 92 then it distributes the master clock information from the best clock on its slave side to the remaining subnet when determining the best master on the remaining subnet. For example, assume that the best clock on the subnet of the first side of the boundary node 92 is the local clock 30 in the node 20. The boundary clock 94 as a consequence of this assumes the role of slave clock to the local clock 30 and adjusts its time to synchronize to a time value distributed using the timing data packet 50 and a corresponding follow up packet on the communication link 12. Under such conditions, when the boundary node 92 transfers a timing data packet via the communication link 90 it includes the master clock information 52 from the node 20 in the timing data packet. As a consequence, the boundary clock 94 may assume the role of master clock on the second side of the boundary node 92 based on the clock information associated with the local clock 30 rather than the boundary clock information 96. This enables the boundary clock 94 to inherit the characteristics of the best master clock to which it is a slave on one subnet for the purposes of evaluating its suitability as a master clock on another subnet.

**[0035]** Figure 4 shows an arrangement for determining a delay associated with the transfer of timing data packets between a pair of nodes 120-122. The delay includes the propagation delay on a pair of communication links 110-112 and the delay associated with an intervening communication device 114. The communication device 114 in one embodiment is a repeater and in other embodiments may be a gateway or router or switching hub or similar device. The arrangement shown enables a correction for this delay and an improvement in the accuracy of time synchronization between a pair corresponding local clocks 124-126.

**[0036]** It is assumed that the node 120 is the master clock and the node 122 is the slave clock and that the delay is symmetrical between the nodes 120-122. Initially, the normal synchronization protocol which is described in U.S. Patent No. 5,566,180 is run to reach syntonization between the local clocks 124-126. Syntonization means that the local clocks 124-126 run at

substantially the same speed although they may not hold the same time. Accordingly, the nodes 120-122 include time packet recognizers that obtain local time values from the local clocks 124-126, respectively, when timing data packets are transmitted and received. The nodes 120-122 also include means for generating timing data packets and follow up packets and processing means for computing differences between time-stamps carried in follow up packets and time values obtained from the local clocks 124-126 and means for adjusting the local clocks 124-126 in response to the computed differences.

**[0037]** The node 120 generates a timing data packet 130 and transfers it via the communication link 110. The node 120 then generates a follow up packet 132 that carries a time-stamp 134. The time-stamp 134 is a time value sampled from the local clock 124 when the timing data packet 130 was generated. The timing data packet 130 when received via the communication link 112 causes the node 122 to sample a time value from the local clock 126. The node 122 then determines an apparent delay ( $D_1$ ) from the node 120 to the node 122 by computing a difference between the time value sampled from the local clock 126 and the time-stamp 134.

**[0038]** Thereafter, the node 122 transfers a delay packet 136 and then a follow up packet 138 via the communication link 112. The follow up packet 138 carries a time-stamp 140 which is a time value sampled from the local clock 126 when the delay packet 136 was generated. In one embodiment, the delay packet 136 is a timing data packet with a special delay flag set. The delay flag indicates to other nodes that the delay packet 136 should be ignored by slave clocks and should not be used to determine a best master clock.

**[0039]** The delay packet 136 when received via the communication link 110 causes the node 120 to sample a time value from the local clock 124. The node 120 then determines an apparent delay ( $D_2$ ) from the node 122 to the node 120 by computing a difference between the time value sampled from the local clock 124 when the delay packet 136 was received and the time-stamp 140.

**[0040]** Thereafter, the node 120 transfers a calibration packet 142 via the communication link 110. The calibration packet 142 carries a delay 144 equal to the apparent delay  $D_2$ . The node 122 receives the calibration packet 142 via the communication link 112 and determines the overall delay between the nodes 120 and 122 by computing an average of  $D_1$  and  $D_2$ . The overall delay may then be used by the node 122 as a slave clock when it adjusts the local clock 126 to correct for the delay between the node 120 and 122.

**[0041]** In one embodiment, the nodes 120-122 perform the calibration procedure set forth above on a periodic basis to correct for time variation in the delay. The calibration procedure in different slave clocks may be staggered so that all slave clocks don't try to calibrate at the same time and cause excessive network traffic.

Slave clocks may be staggered using a random number generator to start the first calibration sequence after a node reset.

**[0042]** Figure 5 shows an embodiment of the communication device 114 which includes mechanisms for determining and reporting the amount of jitter introduced by the communication device 14. The arrangement shown enables a correction for the reported jitter and an improvement in the accuracy of time synchronization in between the local clocks 124-126.

**[0043]** The communication device 14 includes a set of physical interface (PHY) circuits 210-214 that interface to the communication links 110-112 and additional communication links such as a communication link 252. The PHY circuits 210-214 recover data and clock signals from transmissions received via the communication links 110-112 and 252 and provide the recovered data and clock signals to a set of start of frame detectors 202-206, respectively, and to repeater circuitry 230.

**[0044]** One of the start of frame detectors 202-206 detects the start of an incoming packet on a corresponding one of the communication links 110-112 and 252 and sets a corresponding flip flop 220-224. The set one of the flip flops 220-224 starts a counter 240 through an or gate 234. The remaining start of frame detectors 202-206 detect the start of outgoing packets on the corresponding communication links 110-112 and 252 and set corresponding flip flops 220-224. When all of the flip flops 220-224 are set the counter 240 is stopped using an and gate 232. Thereafter the counter 240 holds a count that indicates a difference between the time of entry of a packet to the communication device 114 and the time that the last repeated packet is transferred out of the communication device 114.

**[0045]** For example, the start of frame detector 202 detects the start of the timing data packet 130 which is received via the communication link 110 and sets the flip flop 220 which starts the counter 240. The repeater circuitry 230 retransmits the timing data packet 130 via the communication links 112 and 252 using the PHY circuits 212 and 214. The start of frame detectors 204-206 detect the start of outgoing timing data packets 130 on the communication links 112 and 252 and set the flip flops 222-224. The last of the flip flops 222-224 that is set by the outgoing timing data packets 130 stops the counter 240.

**[0046]** The repeater 114 includes a microprocessor 242 that may perform statistical measurements on counts obtained from the counter 240. The microprocessor 242 has the capability of communication with nodes served by the repeater 114. The microprocessor 242 provides transmit data targeted for the nodes and obtains data received from the nodes via signal lines 250. The microprocessor 242 may provide results of its statistical measurements to nodes served by the repeater 114.

**[0047]** The foregoing detailed description of the present invention is provided for the purposes of illustra-

tion and is not intended to be exhaustive or to limit the invention to the precise embodiment disclosed. Accordingly, the scope of the present invention is defined by the appended claims.

## Claims

### 1. A distributed system, comprising:

first node coupled to a communication link (12), the first node having a local clock and a set of information pertaining to the local clock, the first node transferring a packet (50) on the communication link (12) that carries the information (52);

second node coupled to the communication link (12), the second node having a local clock and a set of information pertaining to the local clock of the second node, the second node receiving the packet (50) on the communication link (12) and determining whether the local clock of the second node is a master clock that synchronizes a time value in the local clock of the first node or a slave clock that synchronizes to a time value from the local clock in the first node by comparing the information (52) in the packet (50) to the information pertaining to the local clock in the second node.

2. The distributed system of claim 1, wherein the information pertaining to the local clock includes an indication of whether the local clock is traceable to a standard clock.
3. The distributed system of claim 1, wherein the information pertaining to the local clock includes an indication of whether the local clock appears to be functional and within specification.
4. The distributed system of claim 1, wherein the information pertaining to the local clock includes an indication of whether the local clock is traceable to a standard clock.
5. The distributed system of claim 1, wherein the information pertaining to the local clock includes an estimate of an accuracy associated with the local clock.
6. The distributed system of claim 1, wherein the information pertaining to the local clock includes a unique identifier associated with the local clock.
7. The distributed system of claim 1, wherein the information pertaining to the local clock includes an indication of whether the local clock is a boundary clock in a boundary node that links together a set of sub-nets.

8. The distributed system of claim 1, wherein the second node compares the information (52) in the packet (50) to the local information in the second node using a prioritized set of rules for evaluating the information.
9. The distributed system of claim 1, wherein the second node is the master clock if the packet is not detected on the communication link (12) during a predetermined timeout interval.
10. The distributed system of claim 1, wherein the second node initially ignores a predetermined number of packets that carry the information pertaining to the local clock in the first node.
11. The distributed system of claim 1, further comprising:
 

third node coupled to a second communication link (90), the third node having a local clock and a set of information pertaining to the local clock of the third node, the third node transferring a packet on the second communication link (90) that carries the information;

boundary node (92) coupled to the communication link (12) and the second communication link (90), the boundary node (92) having a boundary clock (94) and a set of information (96) pertaining to the boundary clock (94), the boundary node (92) receiving the packet on the communication link (12) and determining whether the boundary clock (94) is the master clock that synchronizes the time value in the local clock of the first node or another slave clock that synchronizes to the time value from the local clock in the first node by comparing the information in the packet to the information (96) pertaining to the boundary clock (94), the boundary node (92) receiving the packet on the second communication link (90) and determining whether the boundary clock (94) is a master clock that synchronizes a time value in the local clock of the third node or a slave clock that synchronizes to a time value from the local clock in the third node by comparing the information in the packet transferred via the second communication link (90) to the information (96) pertaining to the boundary clock (94).
12. The distributed system of claim 11, wherein the boundary node (92) is the master clock on the communication link (12) if the packet is not detected on the communication link (12) during a predetermined timeout interval.
13. The distributed system of claim 11, wherein the boundary node (92) is the master clock on the sec-

ond communication link (90) if the packet is not detected on the first communication link (12) during a predetermined timeout interval.

14. The distributed system of claim 11, wherein the information (96) pertaining to the boundary clock (94) is derived from the master clock on the communication link (12).
15. The distributed system of claim 11, wherein the information (96) pertaining to the boundary clock (94) is derived from the master clock on the second communication link (90).
16. A distributed system, comprising:
 

communication device (14) that enables communication between a first communication link (110) and a second communication link (112);

slave node that transfers a delay packet (136) and a follow up packet (138) via the second communication link (112), the follow up packet (138) containing a time-stamp (140) that indicates a local time in the slave node at which the delay packet (136) was transferred;

master node that receives the delay packet (136) and the follow up packet (138) via the first communication link (110) through the communication device (14), obtains a time value from a local clock in the master node when the delay packet (136) is received and then uses the time-stamp (140) and the time value to determine an apparent delay (144) in the transfer of the delay packet (136) from the slave node to the master node, the master node then transferring the apparent delay (144) to the slave node in a calibration packet (142) such that the slave node then uses the apparent delay (144) to determine a delay in the transfer of a timing data packet (130) between the master and the slave nodes.
17. The distributed system of claim 16, wherein the slave node determines the delay by computing an average of the apparent delay (144) in the transfer of the delay packet (136) from the slave node to the master node and an apparent delay (144) in the transfer of a timing data packet from the master node to the slave node.
18. The distributed system of claim 17, wherein the slave node determines the apparent delay (144) in the transfer of the timing data packet (130) from the master node to the slave node by receiving the timing data packet (130) and obtaining a second time value from a local clock in the slave node when the timing data packet (130) is received, and the receiving a follow up packet (138) corresponding to the

timing data packet (130) from the master node which contains a second time-stamp that indicates a local time in the master node at which the timing data packet (130) was transferred and then using the second time stamp and the second time value to compute the apparent delay (144) in the transfer of the timing data packet (130) from the master node to the slave node.

19. The distributed system of claim 16, wherein the delay packet (136) is a timing data packet with a flag that indicates the delay packet (136) such that a set of additional nodes coupled to the first and second communication links (110-112) do not attempt to use the timing data packet with the flag for selection of master and slave clocks.
20. The distributed system of claim 16, wherein the communication device (14) includes means for determining a synchronization error between the local clock in the master node and a local clock in the slave node which is caused by jitter associated with the communication device (14).
21. The distributed system of claim 20, wherein the communication device (14) further includes means for communicating the synchronization error to the master and slave nodes via the first and second communication links (110-112).



FIG. 1

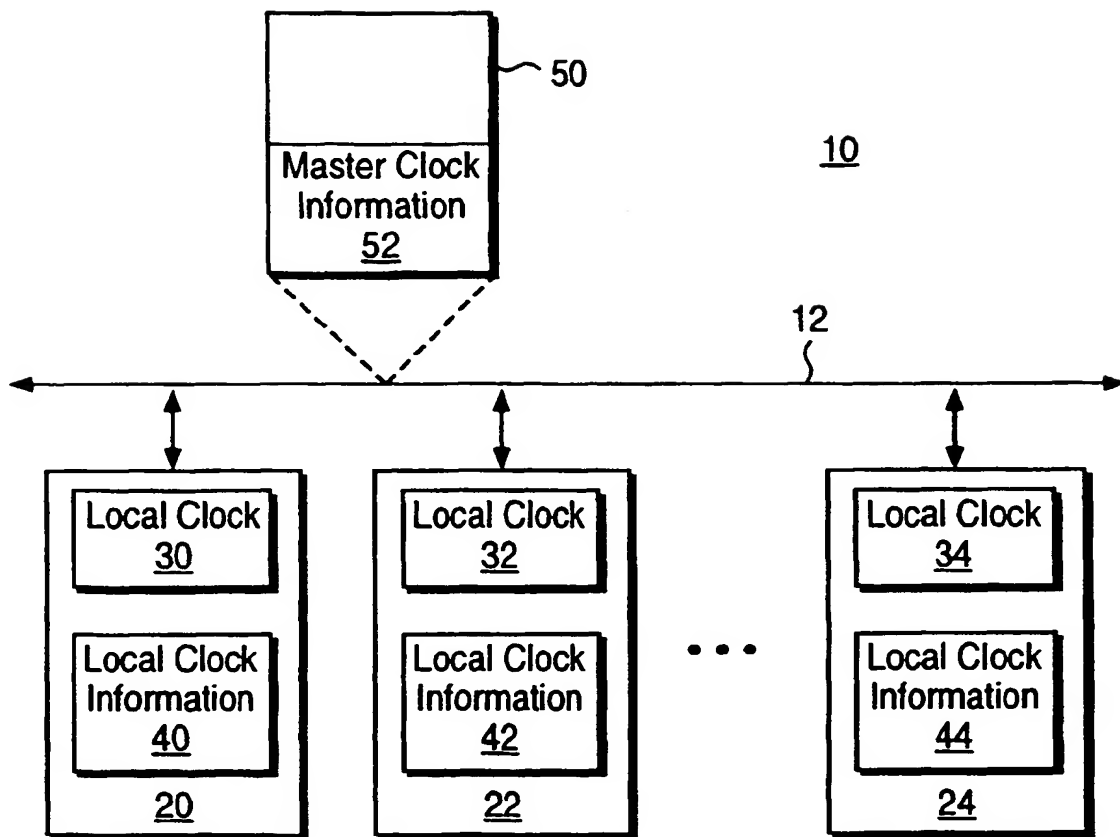


FIG. 2

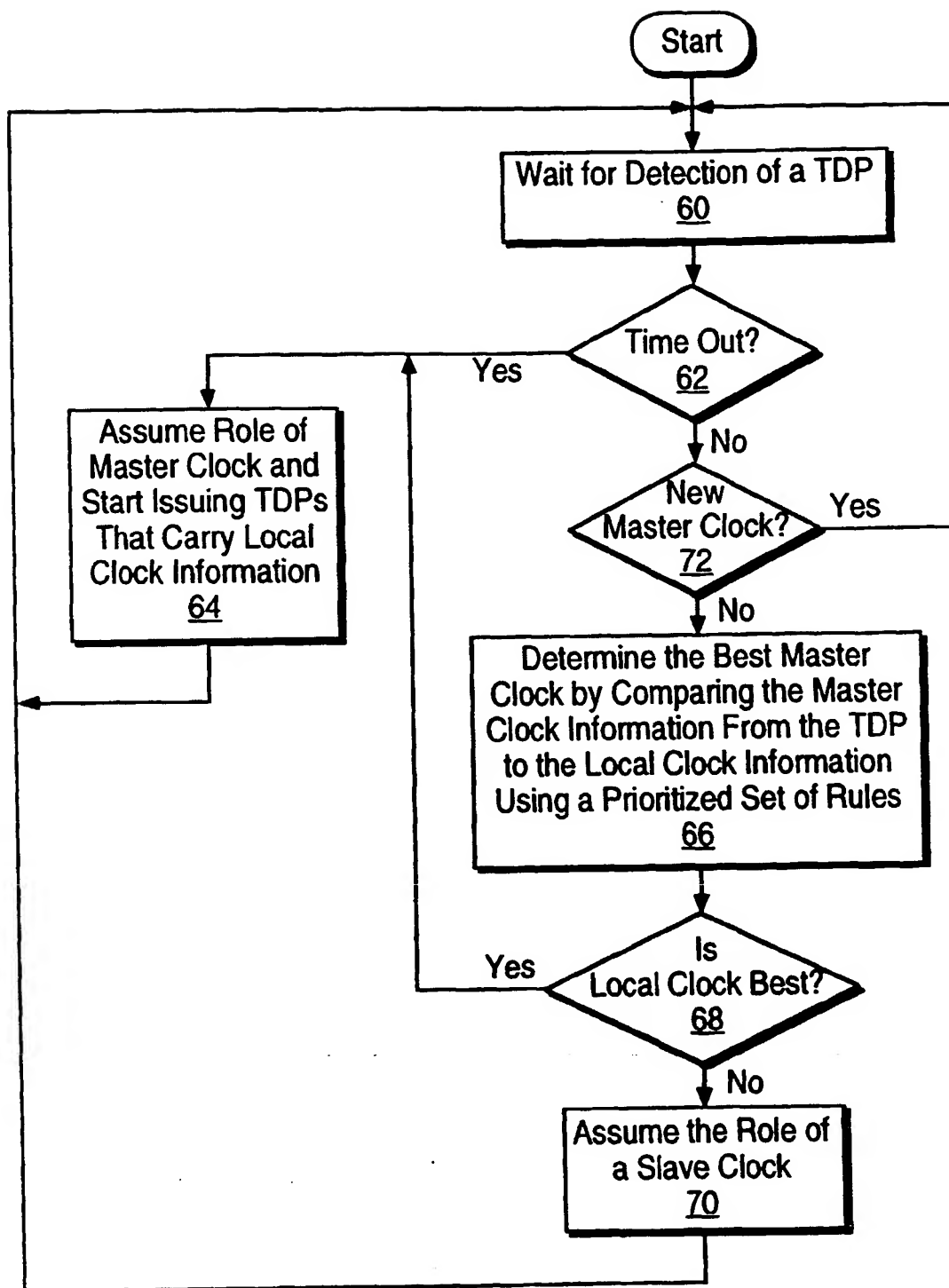


FIG. 3

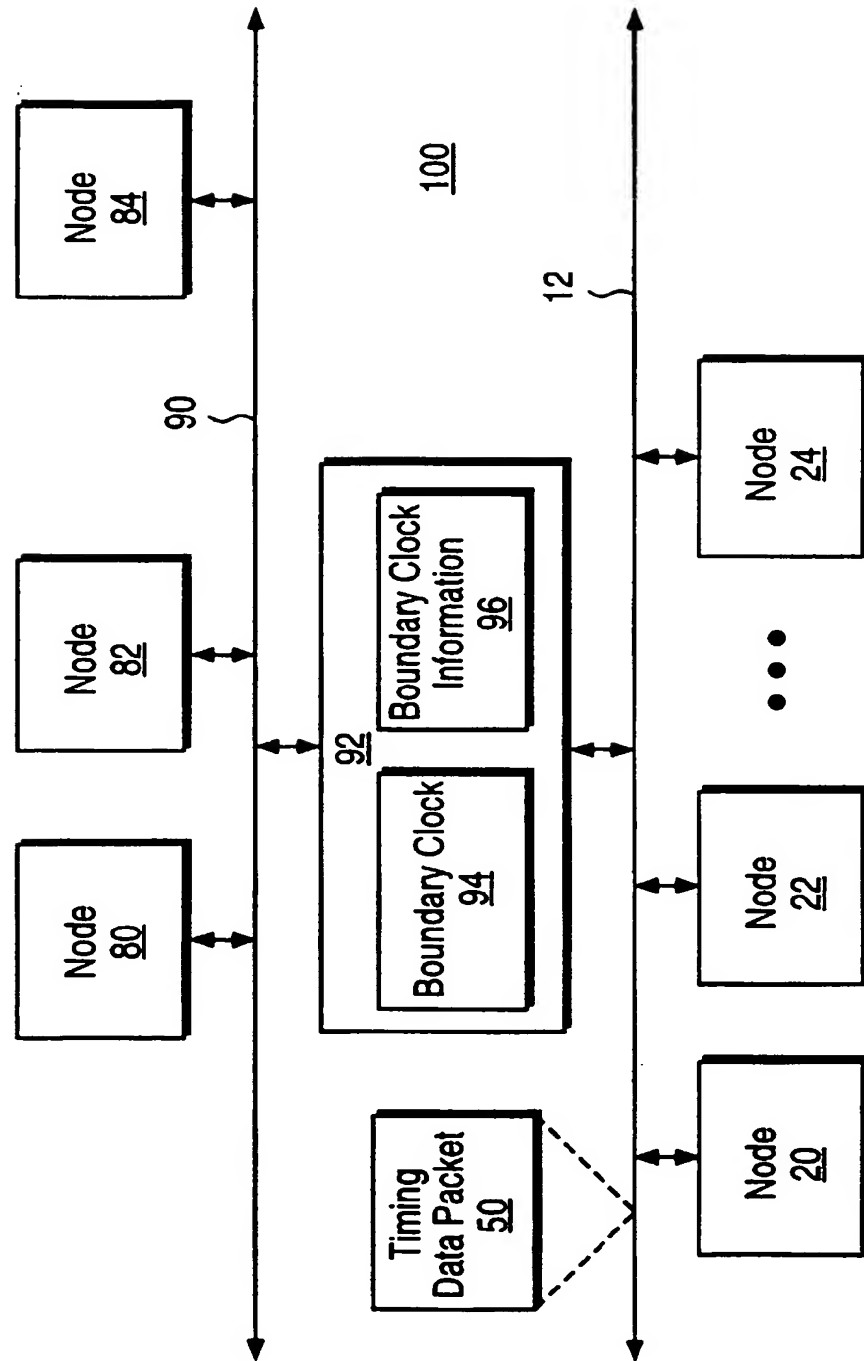


FIG. 4

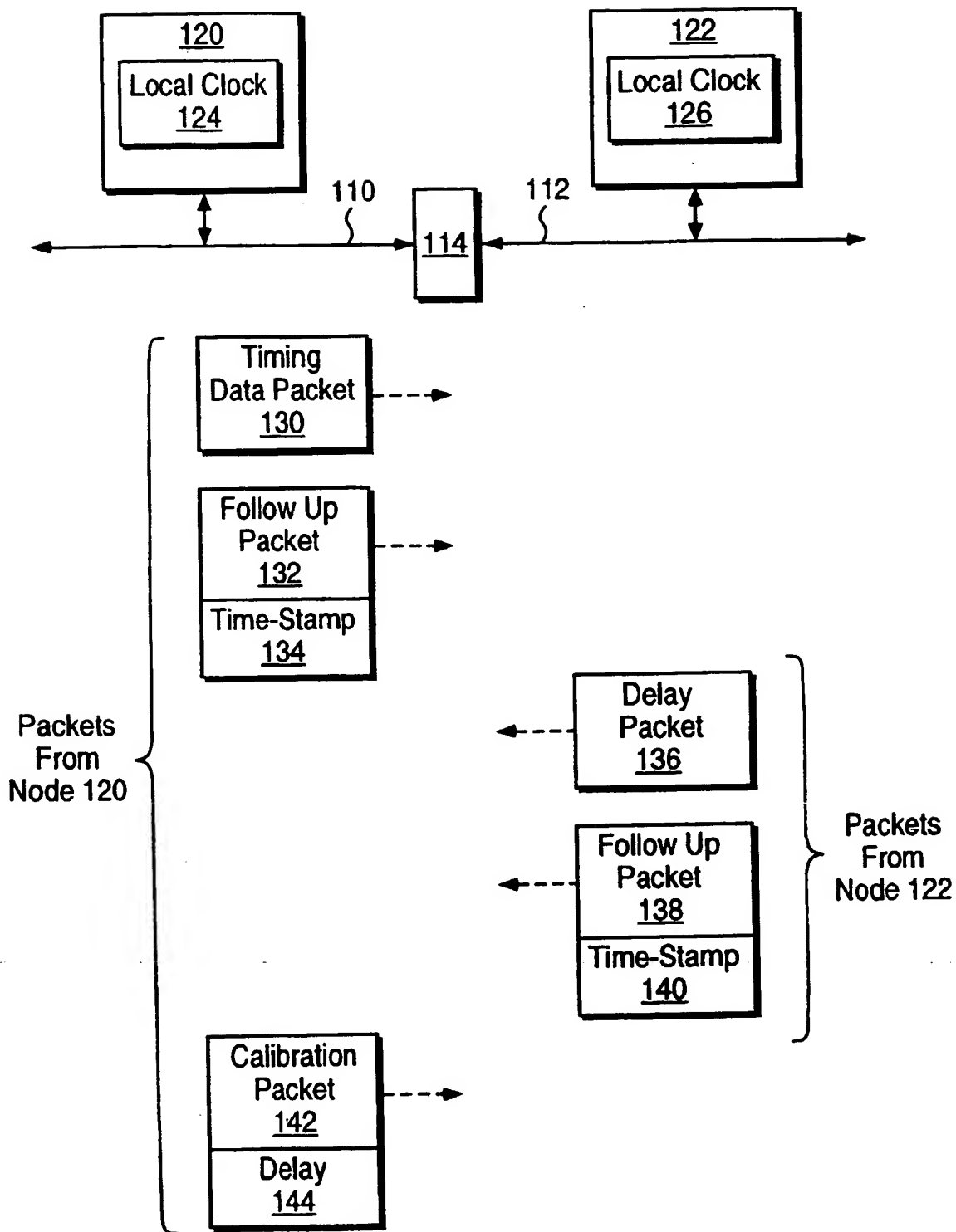
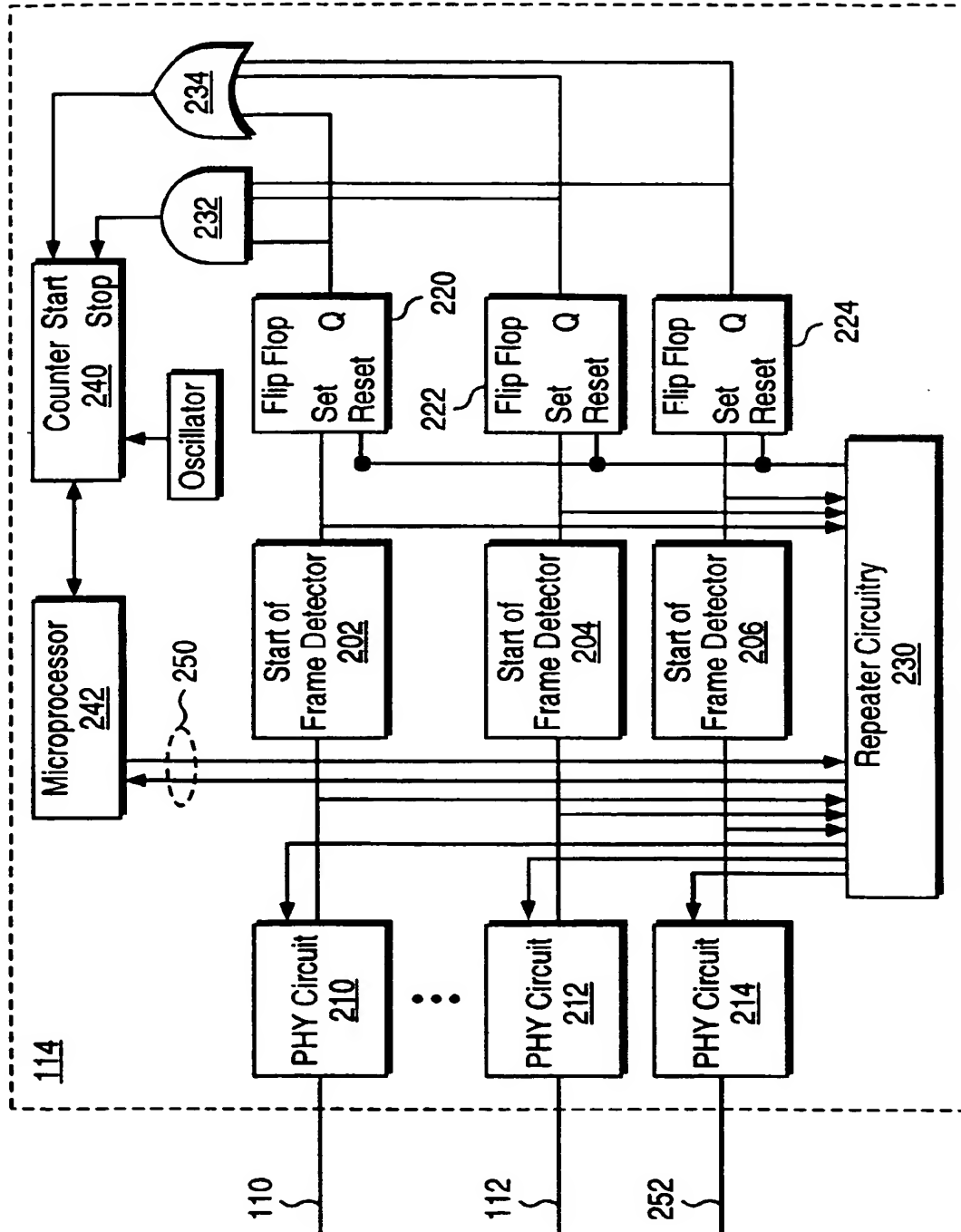
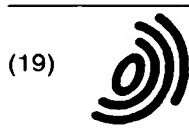


FIG. 5



**THIS PAGE BLANK (USPTO)**



(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:  
06.05.2004 Bulletin 2004/19

(51) Int Cl.7: H04J 3/06

(43) Date of publication A2:  
05.04.2000 Bulletin 2000/14

(21) Application number: 99111373.9

(22) Date of filing: 10.06.1999

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE  
Designated Extension States:  
AL LT LV MK RO SI

(72) Inventor: Eidson, John C.  
Palo Alto, CA 94303 (US)

(74) Representative: Schoppe, Fritz, Dipl.-Ing.  
Schoppe, Zimmermann, Stöckeler & Zinkler  
Patentanwälte  
Postfach 246  
82043 Pullach bei München (DE)

(30) Priority: 29.09.1998 US 162609

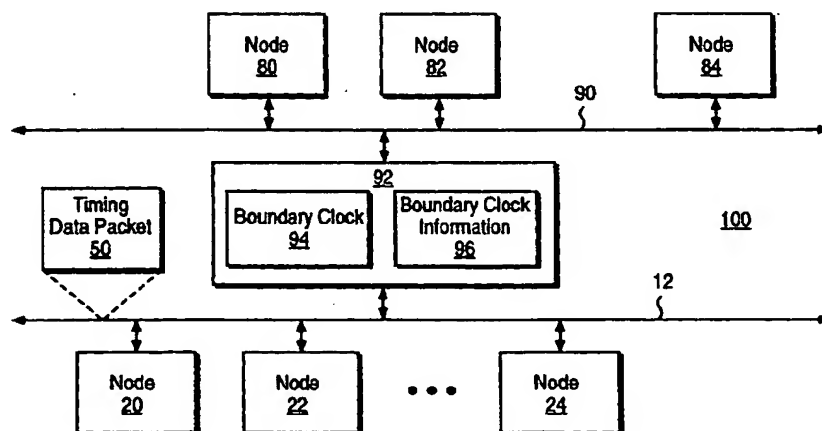
(71) Applicant: Agilent Technologies, Inc. (a Delaware  
corporation)  
Palo Alto, CA 94303 (US)

(54) Organization of time synchronization in a distributed system

(57) A distributed system with mechanisms for automatic selection of master and slave clocks to be used for clock synchronization. The distributed system includes a set of nodes (20-24), including a first node and a second node, each having a local clock (30-34) and a set of information (40-44) pertaining to the local clock (30-34). The first node transfers a packet (50) on a communication link (12) that carries the information (52). The second node receives the packet (50) on the communication link (12) and determines whether the local

clock of the second node is a master clock that synchronizes a time value in the local clock of the first node or a slave clock that synchronizes to a time value from the local clock in the first node by comparing the information (52) in the packet (50) to the information pertaining to the local clock in the second node. Automatic selection of master and slave clocks in boundary nodes (92) is provided along with mechanisms for determining clock synchronization delays and mechanisms for reporting jitter associated with communication devices (14).

FIG. 3





European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 99 11 1373

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.CL7)
X	US 5 519 726 A (DICARLO DAVID A) 21 May 1996 (1996-05-21) * column 5, line 2 - column 9, line 42 * * claims 1,5,6 *	1-15	H04J3/06
X,D	US 5 566 180 A (EIDSON JOHN C ET AL) 15 October 1996 (1996-10-15) * column 4, line 9 - line 16 * * column 6, line 1 - line 19 * * column 7, line 23 - column 8, line 55 * * claims 1-5 *	16-21	
A	EP 0 389 780 A (IBM) 3 October 1990 (1990-10-03) * page 3, line 40 - line 53 * * claim 1 *	1	
A	EP 0 435 395 A (PHILIPS PATENTVERWALTUNG; PHILIPS NV (NL)) 3 July 1991 (1991-07-03) * column 6, line 11 - column 14, line 53 * * claims 6-13 *	1-15	
The present search report has been drawn up for all claims			TECHNICAL FIELDS SEARCHED (Int.CL7)
			H04J
Place of search	Date of completion of the search	Examiner	
MUNICH	5 March 2004	Belloni, P	
CATEGORY OF CITED DOCUMENTS		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons A : technological background O : non-written disclosure P : intermediate document & : member of the same patent family, corresponding document	
X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document			

GPO FORM 1503 03.82 (P04011)





European Patent  
Office

Application Number

EP 99 11 1373

#### CLAIMS INCURRING FEES

The present European patent application comprised at the time of filing more than ten claims.

- ☐ Only part of the claims have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims and for those claims for which claims fees have been paid, namely claim(s):
- ☐ No claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims.

#### LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

see sheet B

- ☒ All further search fees have been paid within the fixed time limit. The present European search report has been drawn up for all claims.
- ☐ As all searchable claims could be searched without effort justifying an additional fee, the Search Division did not invite payment of any additional fee.
- ☐ Only part of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the inventions in respect of which search fees have been paid, namely claims:
- ☐ None of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the invention first mentioned in the claims, namely claims:



European Patent  
Office

**LACK OF UNITY OF INVENTION  
SHEET B**

Application Number  
**EP 99 11 1373**

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

**1. Claims: 1-15**

A distributed system comprising a first node and a second node, both nodes having a local clock and a set of information pertaining to the local clock, wherein the first node transfers a packet carrying the information and the second node receives the packet and determines, by comparing the information in the packet to the information pertaining to the local clock in the second node, whether the local clock of the second node is a master clock that synchronizes a time value in the local clock of the first node or a slave clock that synchronizes to a time value from the local clock in the first node (claim 1).

Claim 11 contains essentially the same features of claim 1 but applied twice to two sub-networks which have in common a boundary node.

**2. Claims: 16-21**

A distributed system comprising a communication device that enables communication between a first communication link and a second communication link, a slave node that transfers a delay packet and a follow up packet via the second communication link, the follow up packet containing a time-stamp that indicates a local time in the slave node at which the delay packet was transferred and a master node that receives the delay packet and the follow up packet via the first communication link through the communication device, obtains a time value from a local clock in the master node when the delay packet is received and then uses the time-stamp and the time value to determine an apparent delay in the transfer of the delay packet from the slave node to the master node, the master node then transferring the apparent delay to the slave node in a calibration packet such that the slave node then uses the apparent delay to determine a delay in the transfer of a timing data packet between the master and the slave nodes (claim 16).

**ANNEX TO THE EUROPEAN SEARCH REPORT  
ON EUROPEAN PATENT APPLICATION NO.**

EP 99 11 1373

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on  
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

05-03-2004

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 5519726	A	21-05-1996	US	5887029 A	23-03-1999
US 5566180	A	15-10-1996	EP	0722233 A2	17-07-1996
			JP	8221336 A	30-08-1996
EP 0389780	A	03-10-1990	US	5001730 A	19-03-1991
			EP	0389780 A2	03-10-1990
			JP	2022683 C	26-02-1996
			JP	2288526 A	28-11-1990
			JP	7058956 B	21-06-1995
EP 0435395	A	03-07-1991	DE	3943052 A1	04-07-1991
			DE	59010197 D1	18-04-1996
			EP	0435395 A2	03-07-1991

EPO FORM 2048

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

**THIS PAGE BLANK (USPTO)**